DOCUMENT RESUME

ED 259 022 TM 850 392

AUTHOR Agnew, E. John
TITLE The Grading Policies and Practices of High School Teachers.
SPONS AGENCY State Univ. of New York, Albany. School of Education.
PUB DATE Mar 85
NOTE 50p.; Also sponsored by the Study of Stanford and the Schools. Paper presented at the Annual Meeting of the American Educational Research Association (69th, Chicago, IL, March 31-April 4, 1985).
PUB TYPE Speeches/Conference Papers (150) -- Reports - Research/Technical (143)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS College School Cooperation; Grades (Scholastic); *Grading; High Schools; Questionnaires; School Policy; School Surveys; *Secondary School Teachers; Statistical Analysis; *Student Evaluation; *Teacher Attitudes; Teacher Characteristics; Teaching Methods; Test Items
IDENTIFIERS Study of Stanford and the Schools

ABSTRACT

This paper examines the grading policies and practices of about 300 teachers from seven high schools in the San Francisco Bay area. Three questions are addressed: (1) Who or what determines grading policies? (2) How do teachers actually calculate grades, that is, what kinds of performances factors are taken into account by teachers? (3) How do teachers evaluate the grading practices and standards at their schools? These questions are reviewed in light of teachers' assigned school, subject area, academic preparation, and primary student referent group. Findings indicate that teachers are the sole arbiters of grading policies, that they award grades in a remarkably consistent manner, and that many are dissatisfied with grading standards. The apparent contradiction between control and satisfaction is discussed, along with several other interesting results. The appendix contains items from the Teacher Questionnaire for the following categories: demographic and sorting items; grading policy items; grading practice items; and grading evaluation items. (Author/DWH)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED259022

THE GRADING POLICIES AND PRACTICES OF HIGH SCHOOL TEACHERS

E. John Agnew

Stanford University

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Agnew, E. John

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

This paper has been prepared for presentation at the 1985 Annual Meeting of the American Educational Research Association, Chicago, Illinois; March 31-April 4, 1985.

This research was supported by the Study of Stanford and the Schools, the Stanford University School of Education, Stanford, California. All rights reserved.

TM 850 392

BEST COPY

TABLE OF CONTENTS

List of Tables and Figures i

Introduction 1

Background Literature on Grading 2

Sample Schools and Teachers,
Questionnaire Items 6

- Schools 6

- Teachers 10

- The Questionnaire 14

Data Analyses 18

- The Determinants of Grading Policy 19

- Grading Practices 22

- Evaluating Student Assessment 30

Findings and Conclusions 34

Appendix A: Questionnaire Items 38

Bibliography 43

THE GRADING POLICIES AND PRACTICES OF HIGH SCHOOL TEACHERS

ABSTRACT

This paper examines the grading policies and practices of about 300 teachers from seven high schools in the San Francisco Bay area. Three questions are addressed: 1) Who or what determines grading policies? 2) How do teachers actually calculate grades, that is, what kinds of performances/factors are taken into account by teachers? 3) How do teachers evaluate the grading practices and standards at their schools? These questions are reviewed in light of teachers' assigned school, subject area, academic preparation, and primary student referent group. Findings indicate that teachers are the sole arbiters of grading policies, that they award grades in a remarkably consistent manner, and that many are dissatisfied with grading standards. The apparent contradiction between control and satisfaction is discussed, along with several other interesting results.

BEST COPY

LIST OF TABLES AND FIGURES

Table 1: Sample School Enrollments by Grade and District Enrollments for 1982-83 7

Table 2: Minority Enrollments in Sample Schools as a Percentage of Total School Enrollment 9

Table 3: Percentile Ranks for the Parent Education Index and 1982 Math, Reading, and Writing Scores on the California Assessment Program's Survey of Basic Skills 9

Table 4: Teacher Characteristics by School 11

Table 5: Highest College Degree Held By Teachers By School 11

Table 6: Principal Teaching Subjects of Teachers By School 12

Table 7: Categories of Teacher Experience 15

Table 8: Categories of Teacher Education 15

Table 9: Ability Groups With Which Teachers Spend Most of Their Time 16

Table 10: Mean Teacher Responses to Items 49d and 49e By Subject Area of Teacher 23

Table 11: Mean Weights (In Percentages) Given to Various Grading Criteria By Teacher Subject Area 28

* * * * *

Figure 1: California High School Enrollment 1982 7

Figure 2: California High School District Enrollment 1982 8

Figure 3: Distribution of Teacher Age 13

Figure 4: Distribution of Years Teaching 13

Figure 5: Who Is Usually Involved in Making Decisions About How to Grade and Evaluate Student Performance? 20

Figure 6: What Usually Determines the Way You Perform the Tasks of Student Grading and Evaluation? 20

BEST COPY

LIST OF TABLES AND FIGURES

(cont'd)

Figure 7: The Importance of Various Referents for the Assignment of Course Grades 23

Figure 8: The Importance of Various Kinds of Student Work for the Assignment of Course Grades 26

Figure 9: The Weights given to Different Aspects of Student Performance for the Assignment of Course Grades 26

BEST COPY

INTRODUCTION

This research was conducted under the auspices of the Study of Stanford and the Schools, a collaborative research effort between the Stanford School of Education and six high school districts in the San Francisco Bay area. The study was initiated three years ago by the President of Stanford, Donald Kennedy, and the Dean of the School of Education, Myron Atkin. The purpose of the Stanford Study is to bring research faculty (from many departments in the university, not just the School of Education) together with local school practitioners to collectively work on the problems facing the schools. The study is in its third year, and many of the findings which have general interest will be published throughout the remainder of 1985 and into the early months of 1986.

The Stanford Study has focused not on a particular set of research questions, but rather on a process of collaboration between the university and the schools. The steps of this process are:

1. Research issues or school concerns are identified by Stanford researchers, school personnel, or individuals from both settings working together.
2. Data are collected and analyzed, usually by researchers and graduate assistants, but in full cooperation with school personnel.
3. Data and analyses are interpreted in discussions with school personnel closest to the concerns/issues under study.
4. Local policy implications are determined - largely by school personnel, but with the aid of researchers.
5. Findings are prepared for publication if they are general enough to warrant wider dissemination.

BEST COPY

The topics which have been the foci in the first three years of the Study have varied substantially. One group has looked at the development of computing curriculums, another has examined the effect of family structure on achievement, yet another is examining student transcripts at several schools in the process of exploring how students move through the high school curriculum.

This research on grading policies and practices was part of a more general examination of teaching as a profession at seven of the participating twelve high schools. The items which form the basis for this paper were culled from a rather large questionnaire which was administered to the teaching staffs at these schools (a more thorough description of the schools and the teacher sample follows). The data on grading were analyzed by the author, and then meetings were held with the teaching staffs at each of the schools in order to discuss the findings. Because some of the findings were remarkably similar and strong across schools it was judged appropriate to prepare this paper for general dissemination.

Background Literature on Grading

All high school teachers must concern themselves with the problem of evaluating student progress. In the vast majority of cases in America today this evaluation culminates in the assignment of a letter grade to represent student progress (achievement) over an academic term. Ninety percent of American high schools reported using letter grades as the principal method for reporting pupil progress in 1977 (ERS, 1977), and this number represented little change over such figures collected over a decade earlier (Pinchak and Breland, 1974).

Much has been written about "letter grades" over the years. Studies of the reliability of grades have been conducted since the early part of the

century (Starch and Elliot, 1912). What was found then has been reinforced repeatedly in subsequent work (Bells, 1930; Tieg, 1952; Bracht, 1967) - the process of grading student work (regardless of subject) is highly subjectivn, unstable across teachers, as well as unstable within teachers over time.

The predictive and differential validity of grades have also been examined in some depth by a number of researchers (Daniel, 1983; Bejar and Blew, 1981, Klugh and Bierly, 1959; Barr, 1961). The findings here are equally consistent. Grades do poorly predicting either college or career success (Bejar and Blew report correlations between high school grades and first year college performance around .45), and the correlations between grades and standardized achievement tests hover around .40 (Daniel, 1983). In fact, grades predict college performance so poorly that one study suggests that it is more efficient from a financial point of view for college admissions personnel to virtually ignore letter grades and rank in class and make decisions based upon other criteria (Miller, 1967).

Besides their unreliability and lack of predictive and differential validity, grades have also been accused of inflicting emotional harm on students. Burton (1983) found that about twenty five percent of a sample of high school teachers believed that grades have a negative effect on students' self concepts.

Other issues which are discussed in the literature are the phenomena of grade inflation, the uses of grades in college admissions formulas, the relationship between grades and aptitude tests, and their use as motivational and disciplinary tools in the classroom.

Despite all the apparent weaknesses of grading, the practice continues with little change from year-to-year. Burton (1983) reported that most

BEST COPY

teachers do not believe grades have a deleterious effect on students. Furthermore, most high school teachers believe that the positive benefits of grades outweigh the negative. Finally, as a significant portion of teachers have indicated, there is simply no realistic alternative available.

The research reported in this paper has a focus somewhat different than that found in most previous studies. Here the actual practice of grading is examined more closely. In addition, the policy and decision making issues regarding grading practices are analyzed. Many of the findings of this research reinforce our intuition about the way grading policies are established and the ways teachers award grades. However, by viewing both policy and practice issues together several interesting questions come to light which may have important implications for the way schools go about the business of assigning letter grades:

==>> The responsibility for determining how to award grades belongs to the individual teacher - yet many teachers are not satisfied with grading standards. Why is there so little satisfaction with standards which are reportedly entirely under their own control?

==>> The data indicate that there are some differences between schools in the amount of influence on grading exercised by principals and department heads. While there appears to be little statistical relationship between principal/department head involvement and satisfaction with grading policies and practices, it is interesting that on a simple rank-order basis, those schools where there is the most administrative involvement, there is also the highest teacher satisfaction with grading standards. Statistical significance, in this case, may be hidden by the form of the questionnaire items and the conservative nature of the tests used.

These and other findings are described in more detail in the remainder

of this paper. The section Sample Schools and Teachers, and the Questionnaire describes the sample of schools and teachers for this research. In addition, the questionnaire items which are the focus of subsequent analyses are described, including the categorical variables used to stratify the teacher group. The section on Data Analyses describes the statistical methods utilized, and presents the results in a straightforward manner with a minimum of interpretation. The section on Findings and Conclusions summarizes the more salient issues and questions which surfaced through this work, and suggests some potential policy implications.

SAMPLE SCHOOLS AND TEACHERS, AND QUESTIONNAIRE ITEMS

Schools:

The districts and individual schools participating in the study were self-selected on the basis of willingness of superintendents, principals, and teachers to participate. The sample is thus non-random in any sense of the word. However, as the demographic data demonstrate, a fair degree of variability is represented within the sample school group.

As was mentioned previously, the teachers at seven of the twelve participating high schools received the Stanford Study Teacher Questionnaire. Table 1 contains the enrollments by district, school, and grade level. Figure 1 displays where the sample schools fall in terms of the distribution of total high school enrollments for the state of California. Figure 2 displays where the sample schools fall in terms of the distribution of district enrollments in the state (distribution contains just unified and high school districts).

Minority enrollments for the seven schools are shown in Table 2. Minority enrollments range from ninety three percent to twenty five percent of the total student body. Of the seven schools, five are predominently white, while two are predominently Asian.

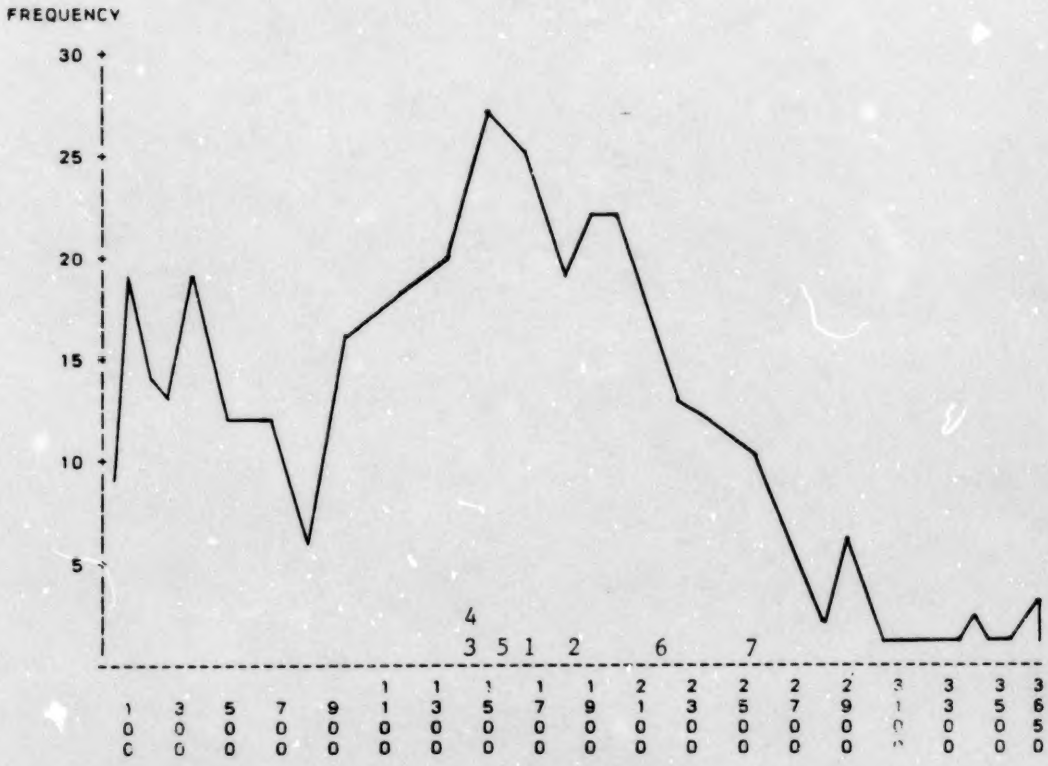
The mean parent education index and percentile rank for the index within the state are displayed in Table 3. Four of the seven schools are at or above the eightieth percentile, while the remaining three schools are below the fiftieth percentile. One school is at the ninth percentile on the parent education index.

The state percentile ranks for the mean reading, writing, and math scores on the California Assessment Program's Survey of Basic Skills are contained in Table 3 as well. Note that only two schools scored at or below

BEST COPY

FIGURE 1

CALIFORNIA HIGH SCHOOL ENROLLMENT
763 HIGH SCHOOLS IN 1982
NUMBERS INDICATE POSITIONS OF SAMPLE SCHOOLS



Source: California Basic Educational Data System, California State Office of Education, Sacramento, CA: 1982

TABLE 1

Sample School Enrollments by Grade

and District Enrollments for 1982-83

		Ninth	Tenth	Eleventh	Twelfth	Total	District
School	1:	0	589	592	511	1,692	8,195
	2:	438	489	511	448	1,887	3,476
	3:	363	367	352	370	1,463	7,765
	4:	0	447	496	478	1,421	32,877
	5:	437	413	395	330	1,580	9,606
	6:	594	666	585	355	2,200	58,371
	7:	682	680	612	551	2,525	58,371

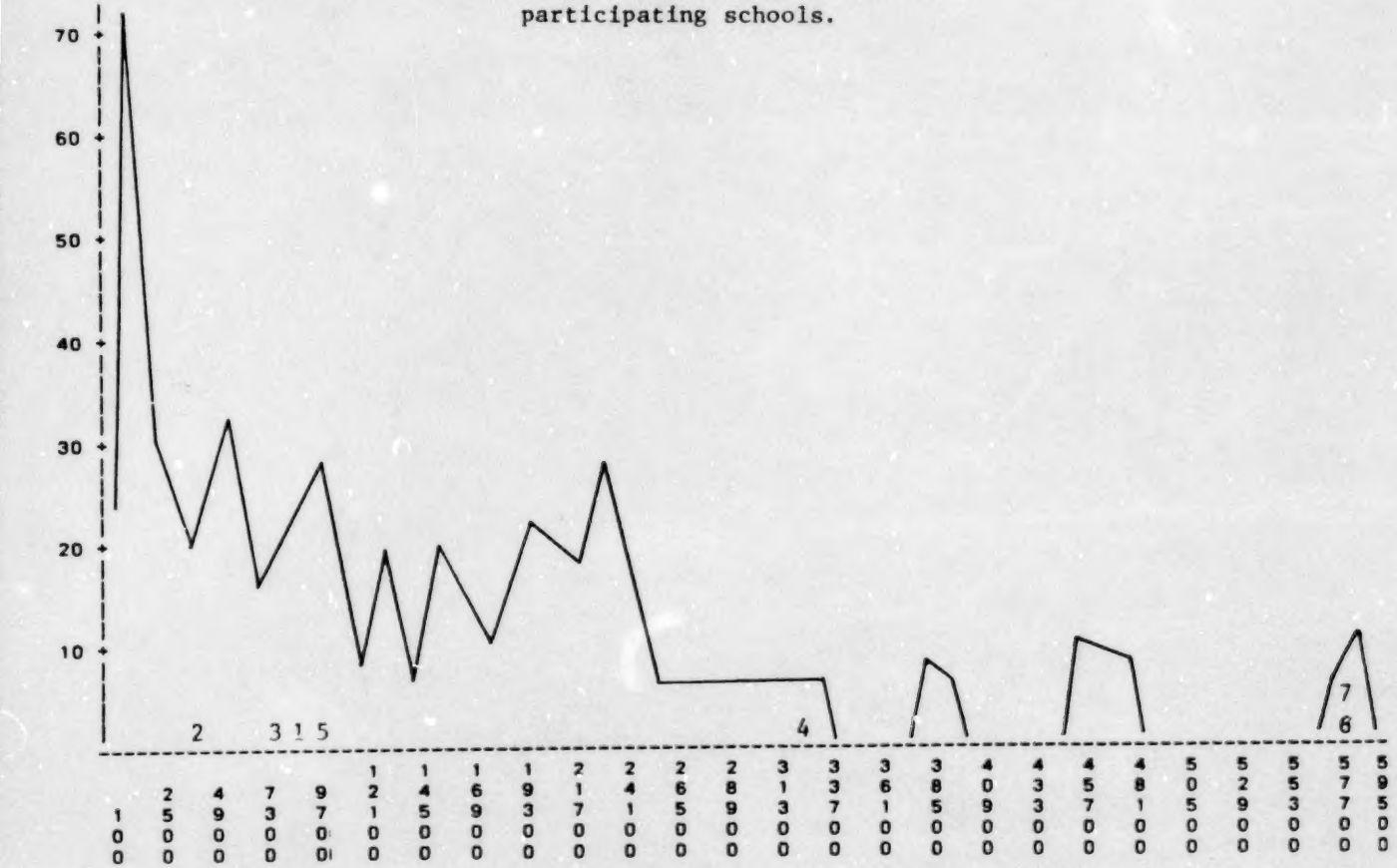
Source: California Basic Educational Data System, California State Department of Education, 1982-83.

FIGURE 2

CALIFORNIA HIGH SCHOOL DISTRICT ENROLLMENT
763 HIGH SCHOOLS IN 1982

Source: California Basic Educational Data System, California State Office of Education, Sacramento.

Note: Numbers indicate positions of participating schools.



Note: This distribution does not include Los Angeles Unified School District, with an enrollment = 530,888, or San Diego City Unified, with an enrollment = 109,442.

TABLE 2

Minority Enrollments in Sample Schools as a Percentage of Total School Enrollment

		Enrollment	%White	%Hispanic	%Black	%Asian	%Filipino	%Am. Indian	Tot Min %
School:	1:	1,692	57.80	15.13	9.69	7.74	1.00	1.00	42.20
	2:	1,887	74.83	8.43	3.34	9.49	3.71	0.21	25.17
	3:	1,463	75.12	6.15	12.58	5.13	0.82	0.21	24.88
	4:	1,421	75.16	18.93	1.34	3.73	0.56	0.28	24.84
	5:	1,582	73.86	8.61	2.28	12.72	2.15	0.38	26.14
	6:	2,200	6.91	2.45	19.14	68.00	3.27	0.23	93.09
	7:	2,525	17.82	11.33	22.57	44.67	3.49	0.12	82.18

Source: California Basic Educational Data System, California State Department of Education, Sacramento, CA; 1983.

TABLE 3

Percentile Ranks[1] for the Parent Education Index[2] and 1982 Math, Reading, and Writing Scores on the California Assessment Program's Survey of Basic Skills[3]

		Parent Education x	Parent Education % Rank	Math % Rank	Reading % Rank	Writing % Rank
School:	1:	2.96	44	84	70	67
	2:	3.95	95	94	91	91
	3:	3.54	85	84	72	79
	4:	3.49	83	86	88	85
	5:	3.46	82	90	80	92
	6:	2.22	9	55	16	26
	7:	2.86	42	76	31	50

1: All percentile ranks are based upon the high schools which particpated in the CAP testing during the 1982-83 school year.

2: Reported by Student. Item: Of your parents, what is the highest educational level reached?

1 = Not a High School Graduate
2 = High School Graduate
3 = Some College
4 = Four-Year College Graduate
5 = Advanced Degree

3: CAP tests administered annually to grades six and twelve. The proportion of seniors who take the test in each school can vary substantially.

Source: California Assessment Program, California Department of Education, Sacramento, CA; 1982.

BEST COPY

the fiftieth percentile on any test scale.

Summary of School Demographic Characteristics: The seven schools from which these data were drawn represent average size schools, from roughly average size school districts in the state of California. Minority enrollments at the seven schools are predominently white or Asian, but at no school is more than seventy six percent of the student body comprised of a single racial or ethnic group. On both achievement and parent education the seven schools represent a wide range, but the distribution is skewed to the right (high) on all four of these measures.

Teachers:

There were roughly 615 teachers at the seven high schools during the spring term of 1983 when this survey was conducted, 389 of which returned completed questionnaires. The questionnaire had several forms. While each of the forms shared a certain number of background items, the remaining portions differed significantly. Forms of the questionnaire were distributed randomly at each school site.

Table 4 contains the sample sizes and descriptive statistics for the characteristics of the teachers at each of the sample schools. This table includes only the 266 teachers who completed the items which are the basis for this paper.

The response rate varies considerably between schools due to several factors. Some schools distributed questionnaires only to teachers, while others included counselors and administrators as well. Also, while the process of distributing the forms of the questionnaire was randomized, the return rate for the various forms was uneven. Finally, some schools did not receive the questionnaire until right before summer break, and at these schools the response rate was notably lower.

BEST COPY

TABLE 4

Teacher Characteristics By School

(Total number of teachers at each school, sample size, sample as per-)
(centage of all teachers, mean and standard deviation of age and years)
(experience, sex.)

School	Total Teachers	Sample N	%N	x Age	s Age	x Years Teaching	s Years Teaching	Males	Females
1	65	30	46	42	7	13	6	22	8
2	90	55	61	49	10	18	8	31	24
3	91	45	49	52	8	22	6	36	9
4	66	49	74	47	9	18	7	29	19
5	73	38	52	50	8	23	7	27	11
6	110	23	21	52	8	20	7	15	8
7	120	26	22	51	8	21	6	15	26
Total Sample:	615	266	43	49	9	19	7	175	90

Note: There were sixteen missing values for Age, twenty eight for Years Teaching, and one for Sex.

TABLE 5

Highest College Degree Held By Teachers By School

Degree		BA	BS	MA	MS	PHD
School	1:	13	5	4	1	0
	2:	23	2	25	2	1
	3:	15	2	23	3	1
	4:	24	3	17	3	0
	5:	10	1	22	4	1
	6:	10	5	5	1	0
	7:	10	1	11	3	0
Sample:		105	19	107	17	3

No Response: 16

BEST COPY

Teachers who completed the form of the teacher questionnaire containing the items on grading were, on the average, about fifty years old, and had about twenty years of teaching experience. Figure 3 displays the distribution of teacher ages, and Figure 4 displays the distribution of years of teaching experience. Almost two thirds of the responding teachers are male (see Table 4).

Table 5 shows the highest degree held by the teachers in the sample. Bachelor of Arts or Sciences degrees were the highest degree attained by 124 teachers, while 113 held Masters degrees in the Arts or Sciences. Only three teachers in the sample held more advanced degrees (Ph.D or Ed.D.).

Table 6 shows the principal teaching subjects of the teachers in the sample. The overall distribution is quite balanced across all subjects. It is not known why this item was unanswered by so many teachers at school six.

Summary of Teacher Demographic Characteristics: The 266 teachers in the sample represent a balanced mix of subject area expertise and academic backgrounds. The average age and experience is high - as it is across the state of California.

TABLE 6

Principal Teaching Subjects of Teachers

By School

School	Missing	ENG	FA	FL	HS	MA	PE	SCI	SS	VB
1:	0	5	0	2	1	4	6	5	4	3
2:	5	11	4	4	2	7	3	5	9	5
3:	4	5	0	2	2	5	6	3	12	4
4:	3	10	4	4	0	4	7	4	6	7
5:	2	6	2	4	0	5	3	3	7	6
6:	10	2	1	0	0	0	1	4	3	2
7:	4	2	4	0	1	2	1	3	2	7
TOTAL SAMPLE: 28		41	11	22	6	27	27	27	43	34

Legend: ENG = English FA = Fine Arts FL = Foreign Language
HS = Health and Safety MA = Math PE = Physical Education
SCI = Science SS = Social Studies VB = Vocational/ Business

BEST COPY 18

FIGURE 3

DISTRIBUTION OF TEACHER AGE

N = 265 Missing = 1

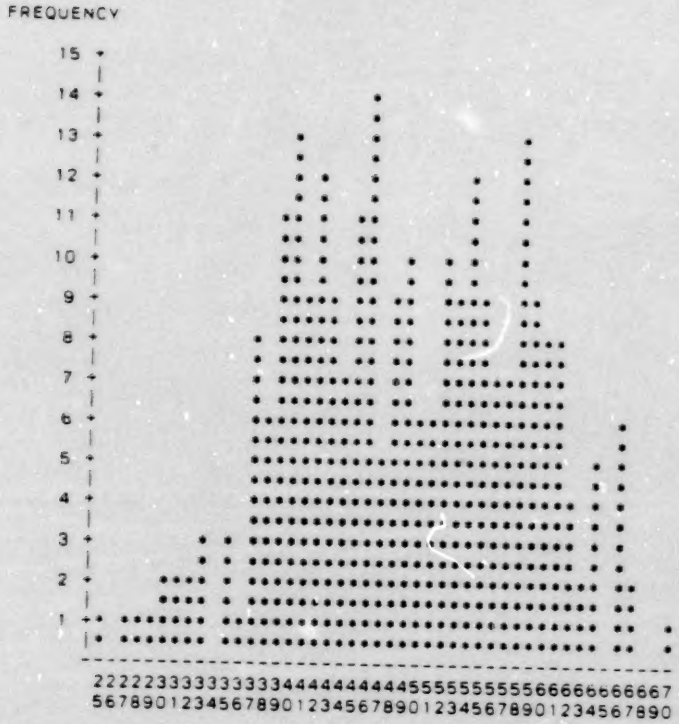
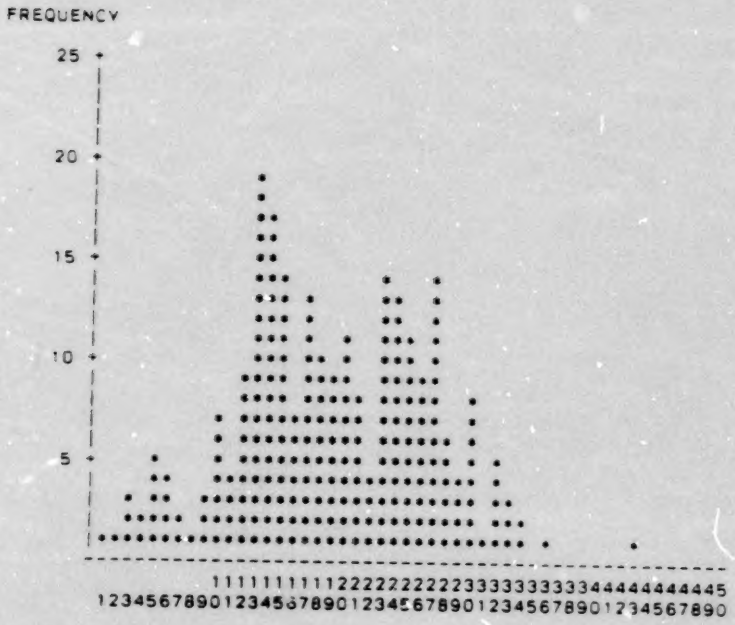


FIGURE 4

DISTRIBUTION OF YEARS TEACHING

N = 246 Missing = 20



BEST COPY

THE QUESTIONNAIRE

The questionnaire itself consisted of seventy items broken up into eight sections: General Background, Professional Activities, Instructional Feedback, Job Satisfaction, School Organization, Students, Tests, and Computers. The full text of the questionnaire is available from the author on request.

The items which are used in this analysis were extracted from various sections: General Background (six items), Job Satisfaction (two items), School Organization (nine items), Students (six items) and Tests (two items). These items are reproduced in full in Appendix A.

Items fall into two groups: 1) Those used for demographic and stratification purposes; and 2) Those pertaining to teachers' grading policies and practices.

Demographic and Stratification Items: The descriptive items used here are assigned school, sex, teaching experience, educational background, principal teaching subject, and the type of student with whom the teacher spent most of his/her instructional day. With the exception of sex, all of these items were used to stratify the teacher sample for comparative analyses.

The breakdowns for assigned school and principal teaching subject were straightforward. The sizes of each of these breakdown groups are displayed in Tables 4 and 6.

The breakdown categories for teaching experience were less than five years, six to ten years, eleven to thirty years, and more than thirty years. The size of each of these groups is displayed in Table 7. It appeared, post hoc, that this many categories were unnecessary as teachers differed very little beyond the fifth year.

Two breakdown categories on educational background were created. The

BEST COPY

TABLE 7

Categories of Teacher Experience

Experience	N	Percent
Less Than 5 Years	39	15
6 to 10 Years	14	5
11 to 30 Years	200	75
30 Years or More	13	5

TABLE 8

Categories of Teacher Education

Category	N	Percent
Breakdown I:		
Bachelors	124	47
Masters +	127	48
Missing	15	6
Breakdown II:		
Liberal Arts/Social Sciences	201	76
Sciences	47	18
Missing	19	7

Note: Assignment in I was on the basis of either having a Bachelor's degree only, or a Master's degree or higher. Assignment to categories in II was based on the degree itself. All those with Bachelors or Masters degrees in the arts alone were grouped, and then those who had either Bachelors or Masters degrees in science were grouped.

BEST COPY

first assigned teachers with Bachelor's degrees (regardless of field) to one group, and teachers with Master's degrees or higher to the other. The second breakdown assigned teachers who had obtained a Bachelor's or Master's in science to one group, and all other teachers to the other. The sizes of these groups are displayed in Table 8.

Finally, the breakdown for type of students taught was made on the basis of teacher responses to item number forty six. Teachers were asked to indicate the percentage of their current students who were high ability, average ability, low ability, and/or remedial (see Appendix A, Item #46). Four breakdown groups were formed on the basis of responses to this item: 1) AVG-HIGH, where seventy five percent of a teacher's students were high or average achieving. 2) AVG-LOW, where seventy five percent of a teachers' students were average or low achieving students. 3) LOW-REM, where sixty five percent of a teacher's students were low achievers or considered remedial students. 4) MIXED, where a teacher could not be assigned to any of the above. This breakdown resulted in mutually exclusive categories. Table 9 displays the sizes of each of the groups.

TABLE 9

Ability Groups With Which Teachers

Spend Most of Their Time

	Ability Group	N	Percent
1.	Average to High	107	40
2.	Average to Low	64	24
3.	Low to Remedial	20	7
4.	Mixed	38	14
	Missing	37	14

Note: Categories are mutually exclusive. Teachers with seventy five percent or more of their students in high or average ability ranges were assigned to 1; teachers with seventy five percent or more of their students in average or low ability ranges were assigned to 2; teachers with sixty five percent or more of their students in low or remedial ranges were assigned to 3; teachers who could not be classified with this scheme were assigned to 4.

BEST COPY 22

Items Pertaining to Grading Policies and Practices: These items focus on three issues: 1) The determinants of grading practices, that is, who exerts influence on grading policies, and what is their degree of influence (five items). 2) How grades are awarded, that is, how teachers actually calculate and award grades to students (3 items). 3) Evaluations of grading and student assessment and promotion procedures in general (6 items). Responses to these items were examined for the whole teacher group, followed by analyses by breakdown groups. These analyses are the focus of the remainder of this paper.

DATA ANALYSES

Data analyses for teacher responses involved two steps: 1) Descriptive statistics were obtained for all content variables for the whole teacher sample and for each of the breakdown groups. 2) Statistical tests were conducted to determine if and where there were differences in teacher responses within breakdown groups. This section will focus on reporting and displaying the results of these analyses according to the three types of content items described previously: Policies, Practices, and Evaluation.

The process used for testing for differences among the breakdown group means involved first a one way anova on the means within each group. If the F was significant ($p < .05$), then the Tukey-Kramer method was used for multiple pairwise comparisons among the means to attempt to locate the differences. Any time the term <u>difference</u> is used in this section of the paper it implies a statistical difference at the alpha = 0.05 level. The Tukey-Kramer method was selected on the basis of its power in situations with unequal cells (SAS, 1982). In one case (items 50a through 50d) a chi-square test of fit was utilized.

Scale means are reported to two decimal places. Following means are the standard deviations in parentheses, e.g. (s=nn). Where the means represent a mean percentage of the teacher sample responding in a certain manner, then the values for both means and standard deviations are rounded to the nearest integer.

The use of these types of tests is not an attempt to lend more legitimacy to these findings than is warranted by the whole nature of the data collection process. It is clear that the best use of the statistics generated in these analyses is to provide a conservative force - preventing undue confidence in conclusions which might be made if the data were inter-

preted on the basis of intuition or experience alone.

Finally, please note that interpretations of the data are kept to a minimum in this section. The majority of interpretation is in the Findings and Conclusions section of the paper.

I. The Determinants of Grading Policy:

To begin with, Item 41d asked teachers to indicate who was involved in decisions regarding how to grade students, and what was their degree of influence. Responses to this item are displayed on Figure 5. Parents exert the least influence with a mean of 1.4 on the four point scale (s=0.68). District Administration is next with a mean of 1.82 (s=0.94), followed by Department Heads with a mean of 2.04 (s=0.96), and Principals with a mean of 2.06 (s=1.01). Teachers indicate that they are clearly the most influential with a mean of 3.86 (s=.047).

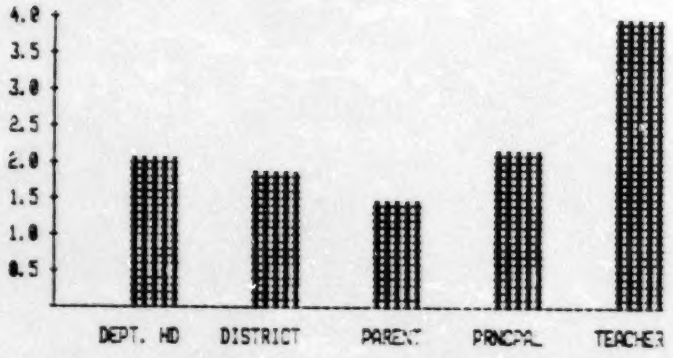
Teachers were also asked to indicate what determined the way they performed the task of grading in Item 39b. Figure 6 displays the responses to this item. Only two percent of the teachers indicated that administrators or department heads determined how grading was done, eleven percent indicated that there were standard operating procedures to follow, thirteen percent indicated that they determined how to grade after consulting with their department head, twenty percent indicated that they determined how to grade after consulting with other teachers, and fifty five percent indicated that they determined how to grade on their own.

Items 43g and 43h further illustrate the low influence of standard operating procedures or school policies on grading practices. Item 43g asked teachers to what extent established school policies determined how grading was done. On the five point scale the mean response was 2.85 (s=1.05), or between "Little" and "Some."

FIGURE 5

Who Is Usually Involved in Making Decisions About How To Grade and Evaluate Student Performance?

Item 41d

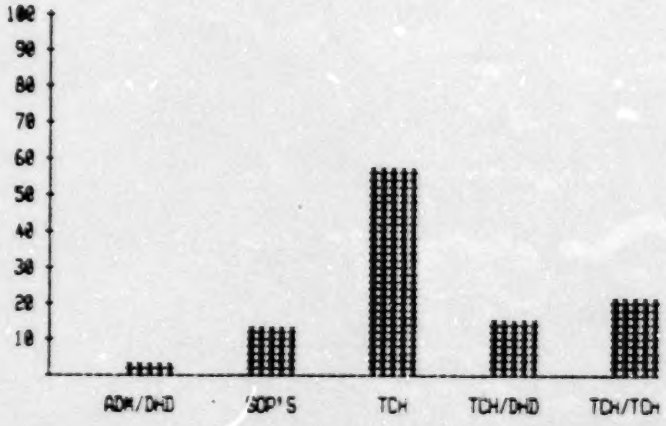


Key: DEPT. HD = Department Heads DISTRICT = District Administration
PARENT = Parents PRINCIPAL = Building Principal
TEACHER = Teacher Alone
Axis Scale: 1 = Not Involved 2 = Little Influence
3 = Moderate Influence 4 = Great Influence

FIGURE 6

What Usually Determines the Way You Perform the Tasks of Student Grading and Evaluation?

Item 39b



Key: ADM/DHD = Building Administrators and/or Department Heads.
SOP'S = There are Standard Operating Procedures to follow.
TCH = Teacher alone decides.
TCH/DHD = Teacher decides after consultation with Department Head.
TCH/TCH = Teacher decides after consultation with colleagues.
Axis Scale: Percentage of sample choosing each response.

Item 43h asked to what extent do establish school policies determine how student progress and achievement is evaluated. On the same scale the mean response was 2.95 (s=.095).

Finally, Item 40b asked teachers to indicate how much freedom they had to determine how to evaluate and grade student achievement. The mean response was 4.55 on a five point scale (s=.69), or between "Considerable" and "Great Deal."

There were few differences on these items among the breakdown groups. There were no differences evidenced on Item 41d on the involvement of the district, teachers, or parents. There were some differences evidenced between schools on the involvement of principals and department heads. These differences will be discussed later in the Findings and Conclusions Section.

There were no striking variations (a chi-square test was used in this instance) between any of the breakdown groups on Item 39b - what usually determines the way grading is done. In every group it appears clear that administrators and department heads are rarely involved, and teachers on their own are the important decision makers. There are small variations between schools in terms of the existence of standard operating procedures, and the frequency of teacher collaboration. Teachers of different subjects exhibit these same minor differences.

There are also no differences among breakdown groups on Items 43g and 43h, both of which relate to the existence of established school policies for grading and student evaluation.

There are statistically significant (at alpha = .05) differences between teachers broken down by school on Item 40b, the freedom to determine how to grade and evaluate student performance. The range of responses between schools on this item is between 4.27 and 4.85 - so the difference

is within a rather narrow range and thus must be interpreted carefully. No other breakdown group exhibited differences on this item.

II. Grading Practices:

Three items on the questionnaire explored how teachers actually award grades. One item dealt with the referents for grades, while the other two focused on the elements of student performance which entered into calculating course grades.

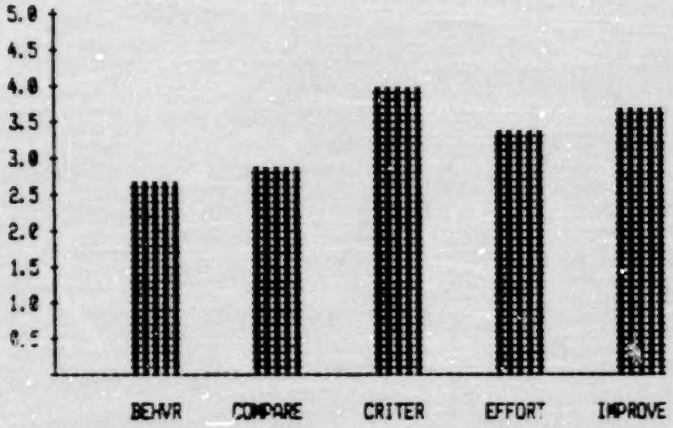
Item 49 asked teachers to indicate how important certain referents were in the assignment of course grades: 1) How individual student work compared with other students' work in the same class (norm-referenced). 2) Whether individual student work met or meets the grade level criteria set for all students (criterion referenced). 3) The amount of improvement the student has shown. 4) How hard the student worked, regardless of the actual level of performance. 5) How well the student behaved in class.

Figure 7 displays the mean responses for all teachers based upon the five points scale: 1. Not Important; 2. Slightly Important; 3. Moderately Important; 4. Very Important; and 5. Extremely Important. Whether student work met or meets grade level criteria is rated the most important with a mean of 3.87 (s=1.07), followed by the amount of improvement with a mean of 3.55 (s=1.01). Next in line was how hard a student worked, regardless of actual level of performance with a mean of 3.30 (s=1.1), followed by comparisons of student work among students with a mean of 2.82 (s=1.23). Rated lowest in overall importance was student behavior in class with a mean of 2.61 (s=1.13).

There were some interesting differences among the breakdown groups on several of these dimensions. The teachers who worked mostly with low and

FIGURE 7

The Importance of Various Referents in the Assignment of Course Grades



Key: BEHVR = Student behavior in class.
COMPARE = How student work compares with others in class.
CRITER = Whether student's work meets grade level criteria.
EFFORT = How hard the student works, regardless of performance.
IMPROVE = The amount of improvement the student has shown.
Axis Scale: 1 = Not Important 2 = Slightly Important
3 = Moderately Important 4 = Very Important
5 = Extremely Important

TABLE 10

Mean Teacher Responses to Items 49d and 49e By Subject area of Teacher

Rate the importance of the following factors in your grading of students at the end of the term:

EFFORT: Subject	Mean	BEHAVIOR: Subject	Mean
Health/Safety	4.00	Fine Arts	3.70
Fine Arts	4.10	P.E.	3.12
P.E.	3.92	Voc/Business	2.93
Social Studies	3.32	Health/Safety	2.80
Voc/Business	3.23	Science	2.70
English	3.22	Social Studies	2.59
Science	3.04	English	2.44
Math	2.93	Foreign Language	1.90
Foreign Language	2.65	Math	1.85

Scale: 1. Not Important 2. Slightly 3. Moderately
4. Very 5. Extremely

BEST COPY

remedial students were the least likely (among the breakdown groups formed by students taught) to compare students to one another in arriving at grades, and they were the most likely to use student class behavior as a factor in grades.

There were no differences among the breakdown groups in terms of the importance of grade level criteria or improvement except between schools, where the range of means went from roughly 2.9 to 4.3, on the importance of grade level criteria, and from about 3.3 to 3.9 on the importance of improvement. School differences were also evidenced on the importance of behavior as a grading criteria. Here school means ranged from about 2.2 to 3.4.

There appears to be no strict pattern to these school differences, except in the case of one school. The means in this school for the importance of improvement, effort, and behavior are higher than every other school, and the mean for the importance of grade-level criteria is lower than every other school. The mean for the importance of within school comparisons is right in the middle of the group. This school has the highest minority enrollment, and the lowest parent education index, and the lowest test scores, of the schools in the sample.

There were two dimensions which evidenced some differences by teacher subject area; the importance of effort, and the importance of behavior. The rank orderings with means for these are provided in Table 10.

As can be seen in Table 10, effort and behavior are less important on average, in the academic subjects, while these criteria are more important in the non-academic subjects.

Item 54a asked teachers to rate the importance of various types of student work in calculating course grades. The scale values were: 1. Not Important; 2. Slightly Important; 3. Moderately Important; and 4. Very

BEST COPY

Important. Figure 8 displays the mean responses on this item for the whole teacher sample. The mean value for worksheets, problem sets, and essays, etc., and for quizzes, was 3.04 (s=.97 for worksheets, s=.85 for quizzes). The mean value for unit tests or mid-terms was 3.65 (s=.66), for final exams the mean was 3.39 (s=.77), and for achievement tests the mean was 2.11 (s=1.19).

School differences appeared in terms of the importance of quizzes. The mean for the school reported above which emphasized effort, improvement, and behavior, was the lowest for all schools on this dimension.

The means of the various subject groups on the importance of worksheets, problem sets, etc., also evidenced some differences. The highest means were for English (highest mean at 3.6, s=.74), Health and Safety, and Social Studies, while the lowest were for Math (lowest mean at 2.48, s=1.05), Foreign Language, and Fine Arts.

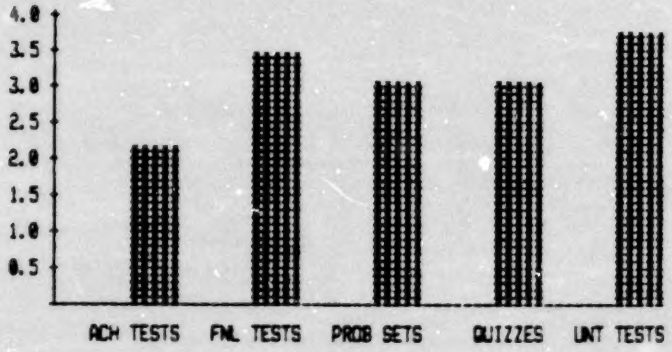
There were differences in terms of the importance of achievement tests among the breakdown groups based on type of students taught. The most striking difference is the relatively high (mean of 2.86, s=1.35) importance of achievement tests for teachers of low and remedial students. This is perhaps due to the increasing emphasis on district or state proficiency tests for the evaluation of marginal students.

There were no differences among any of the groups in terms of the importance of unit tests and midterms, or the importance of final exams.

In Item 59 teachers were asked to weight the various factors which contributed to their course grades. This item attempts to get at the importance of certain criteria, like those of Item 54 which were just discussed, but relative to each other as opposed to an abstract scale. Figure 9 displays the mean weights for the whole teacher sample.

FIGURE 8

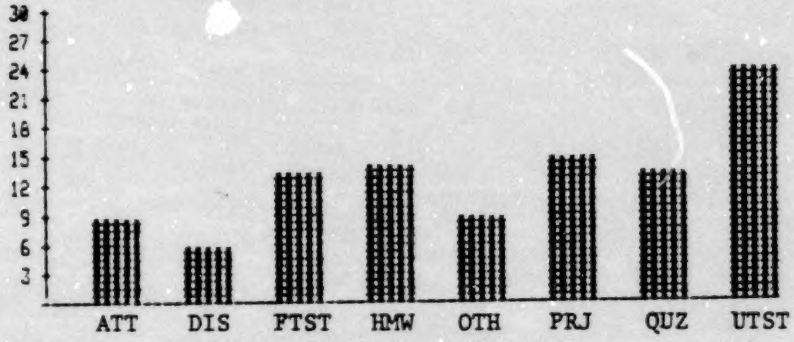
The Importance of Various Kinds of Student Work for the Assignment of Course Grades



Key: ACH TESTS = Achievement tests, like district-wide achievement tests or departmental end-of-course tests. NOT California Assessment Program tests, or minimum proficiencies.
FNL TESTS = Final examinations.
PROB SETS = Worksheets, problem sets, essays, etc.
QUIZZES = Quizzes.
UNT TESTS = Unit tests or mid-terms.
Axis Scale: 1. Not Important 2. Slightly Important
3. Moderately Important 4. Very Important

FIGURE 9

The Weights Given to Different Aspects of Student Performance for the Assignment of Course Grades



Key: ATT = Attendance at class.
DIS = Class discussion.
FTST = Final examinations.
HMW = Homework.
OTH = An aspect not mentioned in this list.
PRJ = Theses, class projects, or similar work.
QUZ = Quiz performance.
UTST = Unit test or midterm performance.
Axis Scale: Percentage weight given to each dimension. Rounding errors and some missing values result in sum less than 100 for these overall mean values.

On the average, homework made up fourteen percent of the grade (s=13), themes, class projects, and the like, also made up fourteen percent of the grade (s=19). Class discussion made up only five percent (s=7), and attendance made up slightly more at eight percent (s=16). Quizzes made up about thirteen percent of the grade (s=14), unit tests and mid-terms made up twenty three percent (s=19), and final exams made up about thirteen percent (s=10).

The category "other" accounted for about eight percent of the grade (s=19). This category was usually chosen by teachers who used performance or skill ratings (in fine arts, vocational/business, or physical education). Science teachers also put lab work in this category. A few teachers used it to weight behavior, effort, and improvement. These cases show up more prominently in the subject-oriented breakdown grouping (see Table 11).

The tests reveal some minor differences on some of these weights for experience and academic background groups. These differences, however, appear so small and inconsistent on examination that they were judged to be spurious. Differences were found between subject groups on the weights given to themes and class projects, class discussion, attendance, and unit tests or midterms. No differences were found for any groups on the weights for homework, quizzes, or final examinations.

The differences between the various subjects are what would be expected based upon even an intuitive sense for the different teaching methods used in the various subject areas. The rather high standard deviations for the mean weights for the whole sample reflect the variation one would expect on this basis. Rather than go through each of these differences in detail, Table 11 contains the weights for each area broken down by subject area.

TABLE 11

Mean Weights (In Percentages) Given to Various Grading Criteria

By Teacher Subject Area (Items 59a-h)

N = 251

VARIABLE	LABEL	N	MEAN	STANDARD DEVIATION
SUBJECT = MATH				
59A	% HOMEWORK IMPORTANT IN GRADING	26	15.62	9.24
59B	% PROJECTS IMPORTANT IN GRADING	26	0.85	2.33
59C	% CLASS DISCUSSION IMPORTANT IN GRADING	26	0.85	1.85
59D	% ATTENDANCE IMPORTANT IN GRADING	26	0.58	2.16
59E	% QUIZZES IMPORTANT IN GRADING	26	19.23	20.23
59F	% UNIT TESTS/MIDTERMS IMPORTANT IN GRADING	26	43.65	24.56
59G	% FINAL EXAM IMPORTANT IN GRADING	26	12.88	9.29
59H	% OTHER IMPORTANT IN GRADING	26	1.92	5.67
SUBJECT = PHYSICAL EDUCATION				
59A	% HOMEWORK IMPORTANT IN GRADING	19	10.00	17.40
59B	% PROJECTS IMPORTANT IN GRADING	19	3.16	6.91
59C	% CLASS DISCUSSION IMPORTANT IN GRADING	19	2.37	5.37
59D	% ATTENDANCE IMPORTANT IN GRADING	19	27.11	24.23
59E	% QUIZZES IMPORTANT IN GRADING	19	6.05	10.75
59F	% UNIT TESTS/MIDTERMS IMPORTANT IN GRADING	19	12.37	22.07
59G	% FINAL EXAM IMPORTANT IN GRADING	19	10.26	13.99
59H	% OTHER IMPORTANT IN GRADING	19	26.05	29.09
SUBJECT = SCIENCE				
59A	% HOMEWORK IMPORTANT IN GRADING	25	13.28	9.12
59B	% PROJECTS IMPORTANT IN GRADING	25	12.12	13.94
59C	% CLASS DISCUSSION IMPORTANT IN GRADING	25	5.48	6.59
59D	% ATTENDANCE IMPORTANT IN GRADING	25	4.04	6.07
59E	% QUIZZES IMPORTANT IN GRADING	25	14.96	16.51
59F	% UNIT TESTS/MIDTERMS IMPORTANT IN GRADING	25	27.72	16.56
59G	% FINAL EXAM IMPORTANT IN GRADING	25	14.20	9.54
59H	% OTHER IMPORTANT IN GRADING	25	7.24	12.73
SUBJECT = SOCIAL STUDIES				
59A	% HOMEWORK IMPORTANT IN GRADING	32	17.63	16.01
59B	% PROJECTS IMPORTANT IN GRADING	32	13.41	8.54
59C	% CLASS DISCUSSION IMPORTANT IN GRADING	32	7.72	8.27
59D	% ATTENDANCE IMPORTANT IN GRADING	32	6.56	10.58
59E	% QUIZZES IMPORTANT IN GRADING	32	11.00	12.42
59F	% UNIT TESTS/MIDTERMS IMPORTANT IN GRADING	32	22.91	16.20
59G	% FINAL EXAM IMPORTANT IN GRADING	32	13.28	9.56
59H	% OTHER IMPORTANT IN GRADING	32	7.97	18.35
SUBJECT = VOCATIONAL/BUSINESS				
59A	% HOMEWORK IMPORTANT IN GRADING	31	12.74	13.71
59B	% PROJECTS IMPORTANT IN GRADING	31	22.26	29.24
59C	% CLASS DISCUSSION IMPORTANT IN GRADING	31	4.39	8.72
59D	% ATTENDANCE IMPORTANT IN GRADING	31	5.35	10.39
59E	% QUIZZES IMPORTANT IN GRADING	31	10.16	10.68
59F	% UNIT TESTS/MIDTERMS IMPORTANT IN GRADING	31	18.65	17.38
59G	% FINAL EXAM IMPORTANT IN GRADING	31	13.55	11.78
59H	% OTHER IMPORTANT IN GRADING	31	8.39	22.34

TABLE 11 (Continued)

Mean Weights (In Percentages) Given to Various Grading Criteria

By Teacher Subject Area (Items 59a-h)

N = 251

VARIABLE	LABEL	N	MEAN	STANDARD DEVIATION
SUBJECT = ENGLISH				
59A	% HOMEWORK IMPORTANT IN GRADING	39	12.38	10.48
59B	% PROJECTS IMPORTANT IN GRADING	39	30.77	17.03
59C	% CLASS DISCUSSION IMPORTANT IN GRADING	39	7.21	7.39
59D	% ATTENDANCE IMPORTANT IN GRADING	39	5.13	9.07
59E	% QUIZZES IMPORTANT IN GRADING	39	11.95	8.96
59F	% UNIT TESTS/MIDTERMS IMPORTANT IN GRADING	39	14.67	9.20
59G	% FINAL EXAM IMPORTANT IN GRADING	39	11.49	7.05
59H	% OTHER IMPORTANT IN GRADING	39	4.87	16.16
SUBJECT = FINE ARTS				
59A	% HOMEWORK IMPORTANT IN GRADING	10	11.00	14.30
59B	% PROJECTS IMPORTANT IN GRADING	10	14.80	22.29
59C	% CLASS DISCUSSION IMPORTANT IN GRADING	10	4.50	6.85
59D	% ATTENDANCE IMPORTANT IN GRADING	10	27.00	28.11
59E	% QUIZZES IMPORTANT IN GRADING	10	9.50	12.57
59F	% UNIT TESTS/MIDTERMS IMPORTANT IN GRADING	10	10.50	11.41
59G	% FINAL EXAM IMPORTANT IN GRADING	10	5.50	6.85
59H	% OTHER IMPORTANT IN GRADING	10	11.50	26.98
SUBJECT = FOREIGN LANGUAGE				
59A	% HOMEWORK IMPORTANT IN GRADING	18	13.61	9.67
59B	% PROJECTS IMPORTANT IN GRADING	18	6.67	7.28
59C	% CLASS DISCUSSION IMPORTANT IN GRADING	18	5.83	8.62
59D	% ATTENDANCE IMPORTANT IN GRADING	18	1.94	5.18
59E	% QUIZZES IMPORTANT IN GRADING	18	18.06	11.65
59F	% UNIT TESTS/MIDTERMS IMPORTANT IN GRADING	18	33.61	14.23
59G	% FINAL EXAM IMPORTANT IN GRADING	18	17.78	12.27
59H	% OTHER IMPORTANT IN GRADING	18	1.11	4.71
SUBJECT = HEALTH AND SAFETY				
59A	% HOMEWORK IMPORTANT IN GRADING	6	20.83	16.86
59B	% PROJECTS IMPORTANT IN GRADING	6	14.17	14.97
59C	% CLASS DISCUSSION IMPORTANT IN GRADING	6	5.83	8.01
59D	% ATTENDANCE IMPORTANT IN GRADING	6	3.33	8.16
59E	% QUIZZES IMPORTANT IN GRADING	6	21.67	18.62
59F	% UNIT TESTS/MIDTERMS IMPORTANT IN GRADING	6	25.00	17.89
59G	% FINAL EXAM IMPORTANT IN GRADING	6	10.00	10.95
59H	% OTHER IMPORTANT IN GRADING	6	0.00	0.00

III. Evaluating Student Assessment Policies and Practices:

Six items focused on how teachers felt about the grading policies and practices at their schools. Items 50a through 50d asked teachers to indicate how they felt about school standards for promotion and failure. Each of these items required teachers to choose among three response options. The items and the responses for the whole teacher sample are reproduced in Table 12.

Item 50b affirms the teachers' comfortableness with the level of their responsibility for setting grading standards and practices. That is, teachers have most of the responsibility for setting these standards (this was discussed earlier under I. Grading Practices), and this is ok with them as a group.

Items 50a and 50b indicate that teachers are in large part unhappy with grading standards for students, both in strictness, and in the amount of responsibility they place on students themselves. This appears to be a contradiction. On the one hand teachers have indicated that they control grading policies and standards, while on the other they seem to be saying that they are not happy with them. This contradiction is discussed further in the Findings and Conclusions section of this paper.

Item 50c is difficult to interpret because of the peculiar phrasing of the response alternatives. Particularly, what are teachers saying when they select the third option - standards are very fair to students? Is this a positive or negative statement? This ambiguity makes this item almost uninterpretable. However, one useful suggestion that has come up as a result of discussions with teachers is to reinterpret the response options as follows:

TABLE 12

Teacher Responses to Items 50a Through 50d

50a: In general, do you feel that the standards for promotion or failure of Students at your school...	
... are too loose?	94 or 37%
... are about right?	156 or 62%
... are too strict?	1 or 0.4%
[15 Missing cases.]	
50b: (Same stem)	
... place too much responsibility on teachers?	35 or 14%
... place about the right amount of responsibility on teachers?	191 or 77%
... place too little responsibility on teachers?	21 or 8%
[19 Missing cases.]	
50c: (Same stem)	
... place too much responsibility on students?	3 or 1%
... place about the right amount of responsibility on students?	108 or 43%
... place too little responsibility on students?	139 or 56%
[16 Missing cases.]	
50d: (Same stem)	
... are generally unfair to students?	10 or 4%
... are moderately fair to students?	83 or 43%
... are very fair to students?	157 or 63%
[16 Missing cases.]	

... unfair to students BECOMES ... biased against students

... moderately fair to students BECOMES ... generally fair to students

... very fair to students BECOMES ... biased in students' favor

To many this reinterpretation makes sense in terms of both the data and their own experience. It is, however, just speculation, and any interpretations of this item must be done with caution.

Two other items were directed at teacher perceptions of the student assessment process. Item 32h asked teachers whether the mechanisms for dealing with failing students [at their school] provided a source of satisfaction for them in their work. Similarly, Item 32i asked teachers whether student evaluation policies were a source of satisfaction. Teachers responded to both items using the following scale: 1. Not At All; 2. Slightly; 3. Moderately; and 4. Very. The mean for Item 32h for all teachers was 2.59 (s=1.05), and for Item 32i it was 2.47 (s=.97). Both values are roughly halfway between "Slightly" and "Moderately."

No differences in the responses to items 50a through 50d could be found using a chi-square test of fit on the cell frequencies of responses crossed with each of the breakdown group categories. There were also no differences in teachers responses to item 32i. Teachers of different subjects did differ on their responses to item 32h - mechanisms for dealing with failing students as a source of job satisfaction. These differences, displayed in Table 13, are perfectly divided between academic and non-academic subjects.

BEST COPY

TABLE 13

Mean Teacher Responses on Item 32h by Subject Area of Teacher

Are School mechanisms for dealing with failing students a source of satisfaction to you in your job?

Non-Academic	|	Health and Safety	3.50	s=0.84
	/	Fine Arts	3.00	s=1.25
	\	Physical Education	2.96	s=0.85
	|	Vocational/Business	2.85	s=1.02
Academic	|	Math	2.63	s=1.11
	|	Foreign Language	2.55	s=1.01
	<	English	2.45	s=1.08
	|	Social Studies	2.32	s=1.04
	|	Science	2.15	s=0.88

Scale: 1. Not At all 2. Slightly 3. Moderately 4. Very

FINDINGS AND CONCLUSIONS

As was stated at the outset, much of what was found in these data confirm our basic intuitions about how grading works in high schools. However, a number of observations, issues and questions arise which are worth looking at and thinking about. The list of findings below are the result of studying these data and talking over the possible interpretations with teachers at each of the participating schools.

1. There is a remarkable degree of consistency to these findings across all schools and all subject areas. While there are some differences, it seems at least a little amazing that there is so much agreement among teachers of all types - particularly on teacher control over grading, and on teacher dissatisfaction with grading.
2. Teachers appear to be the sole arbiters of grading policy. There are school effects demonstrated in a few cases, but the sizes of these differences are small. This does suggest, however, that school administrators (or something at the school site level) can have an impact on grading policy.
3. Teacher subject area appears to have the most influence on how grades will actually be awarded. Statistical differences between teachers were most often found by subject area, and predominantly within those items which dealt with the practice of grading.

   The next most influential categorical variable was the assigned school. Here there were differences in the policy making role of department heads and principals, the freedom teachers were allowed to make these decisions on their own, and in certain types of grading criteria.

   The kind of students a teacher dealt with most of his/her time also

resulted in a few differences. These were entirely within the grading practice items.

Academic background and years experience simply did not appear to make a difference on any dimension.

4. The grading practices of teachers vary in understandable and predictable ways by subject. Striking, however, is the absence of emphasis <u>in any subject whatsoever</u> of class discussion as an important grading criteria. This implies that in the vast majority of high school courses the quality and quantity of student oral expression makes little or no difference to the students' grades. Admittedly class discussion is less relevant in some subjects than in others - but to find it so uniformly unimportant seems indeed a cause for alarm. Perhaps these items did not capture the kinds of oral expression teachers use in their courses? Perhaps, but there were many places where it might have shown up and didn't (the "other" category on item 59 for example). If these findings are not spurious, then to anyone sensitive to the needs of modern life, much less the needs of adolescents for expression, it is clear that this finding identifies a major problem.

5. The next finding is presented with some trepidation due to the small size of the school sample, but the issue posed is too intriguing to pass up. The school in this study which placed the least amount of emphasis on actual learning in the assignment of course grades was also the school with predominantly low-status (measured by parent education) and minority students. Teachers at this school used behavior, attendance, and effort to a much greater degree than grade-level criteria to award grades. At every other school in the sample the situation was the reverse.

The implication of this finding is clear: low status and minority

students are graded not on their academic achievements, but on their attendance and deportment. The value questions this practice raises are self-evident and worthy of public discussion.

As a post-script to this finding, it is interesting to note that of all the schools in the sample, the teachers at this school express the highest degree of dissatisfaction with grading standards and practices.

6. The last finding is perhaps the most intriguing. It was noted in the section on data analyses that while teachers report that their control over grading policies and practices is virtually total, they also report a high degree of dissatisfaction with grading standards. The immediate question this apparent contradiction poses is, why? If the teachers are in control of this aspect of their professional duties, then how can they claim such dissatisfaction - that is, why don't they just exercise their control and change the standards?

In discussions with teachers over this phenomenon two possible explanations come to light:

1) Everybody <u>else</u> has poor standards. The idea here is that the problem rests with other teachers, those who do not have adequately high standards.

2) There is an implicit constraint missing from all of the statements of policy control: Teachers are free to award grades <u>as long as they never fail anyone.</u> This unwritten rule of the grading process appears real enough to most teachers - there are penalties built into the system for failing students, so teachers won't do it. The question is, what is the result of this inability to fail? Certainly one interpretation could go like this: One would predict that for college bound students the effects would be minimal because their

motivation comes from the distinctions between A's, B's, and C's. For the general to low ability student it would seem that the incentive for passing a course is not having to take it again. Thus, severely limiting or eliminating altogether the teacher's power to fail students robs them of one of the only motivational tools available for lower achieving students. Other implications are possible, but in any case the issue is an important one. If there are unwritten rules against failing, then they need to be brought out into the open and examined more carefully by school personnel and the families they serve.

APPENDIX A: TEACHER QUESTIONNAIRE ITEMS

DEMOGRAPHIC AND SORTING ITEMS:

2: School

4: Year of Birth

5: Gender

6: Educational History

7: Career History (teaching only)

12: What do you consider to be your main subject areas?

46: Please estimate the percentage of students in your classes who fall into each academic ability group.

.A % High Achieving

.B % Average Ability

.C % Below Average

.D % Remedial

GRADING POLICY ITEMS:

39b: Think of the way the tasks of grading and evaluating student performance are done at your school. In general, what usually determines how you perform these tasks?

1 = There are standard operating procedures to follow.

2 = School administrators or department heads tell me how to grade and evaluate students.

3 = I decide how to grade and evaluate students after consulting with the department head.

4 = I decide how to grade and evaluate students after consulting with other faculty.

5 = I alone decide the way I will grade and evaluate students.

40b: Under present arrangements, how much freedom do you have to determine how to grade and evaluate student performance?

1 = None 2 = Little 3 = Some 4 = Considerable 5 = Great deal

41d: Who is usually involved in making decisions about how to grade students, and what is their degree of influence?

1= Not Involved 2 = Little Influence 3 = Moderate 4 = Great

.1 District .2 Principal .3 Department Head .4 Teachers

.5 Parents

43g: To what extent do established school policies determine how grading is done?

43h: To what extent do established school policies determine how student progress and achievement is evaluated?

1 = Not at all 2 = Little 3 = Some 4 = Great deal 5 = Completely

BEST COPY

GRADING PRACTICE ITEMS

49: Rate the importance of the following factors in your grading of grading of students at the end of the term:

1 = Not important 2 = Slightly 3 = Moderately 4 = Very 5 = Extremely

.A How student's work compares with that of classmates

.B Whether student's work meets criteria set for all students at that grade level.

.C The amount of improvement the student has shown.

.D How hard the student has worked, regardless of actual level of performance.

.E How well the student behaves in class.

54A: Different types of assignments and tests may be used for different purposes. For each type of assignment or test, rate its importance for the purpose of determining course grades:

1 = Not important 2 = Slightly 3 = Moderately 4 = Very

.1 Worksheets, problem sets, essays, etc.

.2 Quizzes

.3 Unit tests, midterms

.4 Final tests

.5 Achievement tests (Include, for example, district-wide achievement tests or departmental end-of-course proficiency tests. DO NOT include California Assessment Program tests, or minimum proficiency exams.)

BEST COPY

GRADING PRACTICE ITEMS (Cont'd)

59: Please indicate the importance of each source of information to you in assigning grades. Enter a number from 0 - 100 in each blank (percentage) - your numbers should add up to 100%.

.A Homework

.B Theses, class projects, or similar work

.C Class discussion

.D Attendance

.E Quiz performance

.F Unit test or midterm performance

.G Final examination

.H Other (specify)

GRADING EVALUATION ITEMS

32h: Are school mechanisms for dealing with failing students a source of satisfaction to you in your job?

32i: Are your student evaluation policies a source of satisfaction to you in your job?

1 = Not at all 2 = Slightly 3 = Moderately 4 = Very

30: In general, do you feel that the standards for promotion or failure of students in your school...

(Check one in each grouping.)

.A are too loose; are about right; are too strict

.B place too much responsibility on the teacher

place about the right amount of responsibility on the teacher

place too little responsibility on the teacher

.C place too much responsibility on the student

place about the right amount of responsibility on the student

place too little responsibility on the student

.D are generally unfair to the students

are moderately fair to the students

are very fair to the students

B I B L I O G R A P H Y

Barr, A. S. et al. Wisconsin Studies of the Measurement and Prediction of Teacher Effectiveness. Madison: Dembar Publications, 1961.

Bejar, Isaac and Blew, Edwin. "Grade Inflation and the Validity of the Scholastic Aptitude Test." Research Report, College Entrance Examination Board: New York, NY, 1981. ERIC Document Number ED 209 327.

Bells, W. C. "Reliability of Repeated Grading of Essay Type Examinations." Journal of Educational Psychology Vol 21, pp. 48-52; 1930.

Bracht, G. H. "The Comparative Values of Objective Essay Testing in Undergraduate Education: Implications for Valid Assessment of Instruction." Unpublished master's thesis. Denver: University of Colorado, 1967. Cited in Evans, F. B. "What Research Says About Grading." in Simon, S. B. and Bellanca, J. A. Eds. Degrading the Grading Myths. Washington, D.C.: ASCD, 1976. PP. 30-50.

Burton, F. "A Study of the Letter Grade System and Its Effect on The Curriculum." Paper presented at the 1983 annual meeting of the American Educational Research Association, Montreal, Quebec. ERIC Doc No ED 238 143.

Daniel, M. "Correlations of Aptitude Tests with High School Grades: Technical Report, 1983-84." New York: Johnson O'Connor Research Fund, Human Engineering Laboratory, 1983. Eric Doc Num ED 242 753.

ERS: Educational Research Service Report, Reporting Pupil Progress: Policies, Procedures, and Systems. Arlington, VA: Educational Research Service, Inc., 1977.

B I B L I O G R A P H Y (Cont'd)

Klugh, H. and Bierly, R. "The School and College Ability Test and High School Grades as Predictors of Achievement." Educational and Psychological Measurement. Vol. 19, pp. 625-26; 1959.

Miller, S. Measure, Number, and Weight: A Polemical Statement of the College Grading Problem. Ann Arbor, Michigan: University of Michigan, Center for Research on Learning and Teaching, 1967. Cited in: Kirschenbaum, Simon, and Napier. Wad-Ja-Get? The Grading Game in American Education. New York: Hart Publishing Company, Inc. 1971.

Pinchak, B. M. and Breland, H. M. "Grading Practices in American High Schools: National Longitudinal Study of the High School Class of 1972." Education Digest. Vol. 39, pp. 21-23, March, 1974.

Statistical Analysis System, User's Guide: Statistics. Cary, North Carolina: 1982, pp. 169-170.

Starch, D. and Elliott, E. "Reliability of the Grading of High School Work in English." School Review. Vol. 20, pp. 442-57; idem, "Reliability of Grading Work in Mathematics," School Review. Vol. 21, pp. 254-295; 1913, and "Reliability of Grading Work in History," School Review Vol 21, pp. 676-81; 1913.

Tieg, E. W. Educational Diagnosis. Bulletin No. 18. Monterey, California: California Testing Bureau, 1952. Cited in Kirschenbaum, Simon, and Napier, Wad-Ja-Get? The Grading Game in American Education New York: Hart Publishing, 1971.

BEST COPY